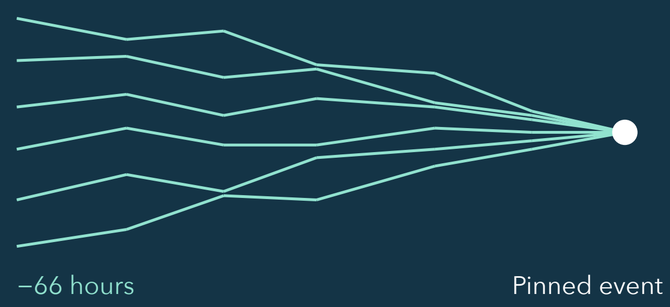
TACKLING CLIMATE CHANGE WITH MACHINE LEARNING / NEURIPS 2026

Run it Back

Autoregression vs. Video Diffusion for Final-Conditioned Weather Trajectories

Jerry Lin and **Elizabeth A. Barnes** · Boston University

Presented by Zeyuan Hu, NVIDIA



Which weather patterns could lead to the same extreme?

Each model samples 1,000 possible 66-hour histories per event, conditioned on the final weather state from ERA5 reanalysis.

REWIND-AR

A probabilistic FourCastNet 3 model starts at the exact ERA5 state and steps back six hours at a time on a roughly 25 km grid.

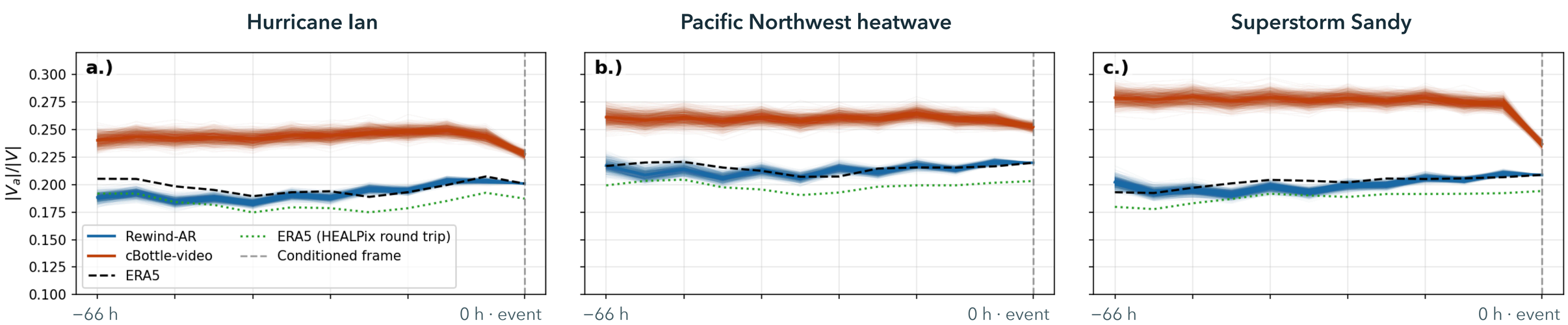
CBOTTLE-VIDEO

A video diffusion model samples all 12 frames together on a roughly 100 km grid. It reproduces the endpoint approximately.

PHYSICAL CONSISTENCY / THREE EVENTS

Rewind-AR stays close to ERA5's midlatitude wind balance

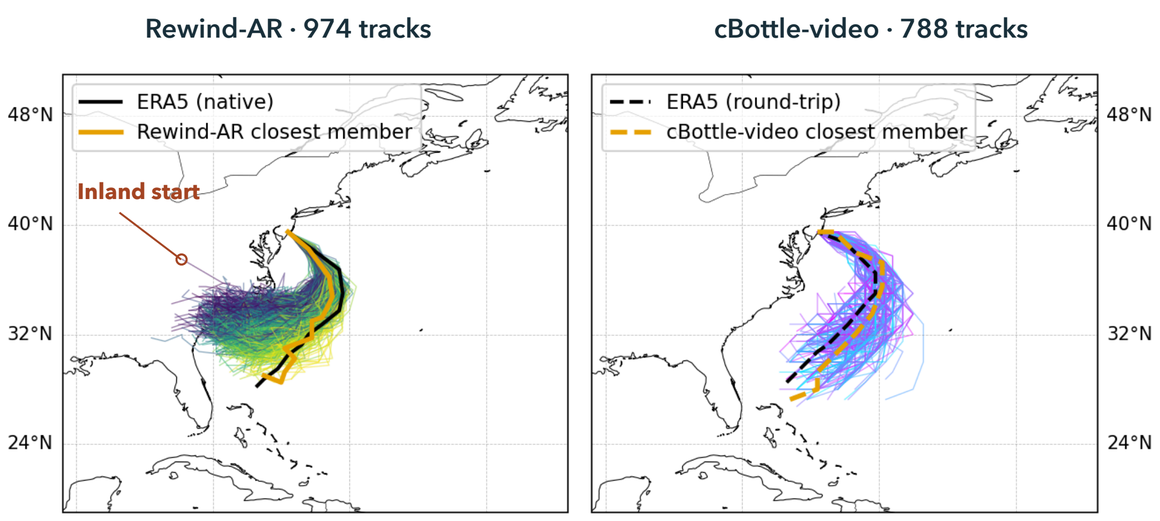
At 500 hPa, we subtract the geostrophic wind inferred from height gradients from the wind at each grid point. The ageostrophic fraction is the ratio of mean residual speed to mean wind speed over 30–60° in both hemispheres.



Thin lines show members and bold lines show ensemble means. Averaged over the 11 frames before each event, Rewind-AR's fraction is **0.002–0.005 below** native ERA5. cBottle-video's is **0.06–0.09 above** its grid-matched ERA5 reference. Rewind-AR's winds still average 1.1–1.8 m/s too weak at −66 h.

SUPERSTORM SANDY / OCTOBER 2012

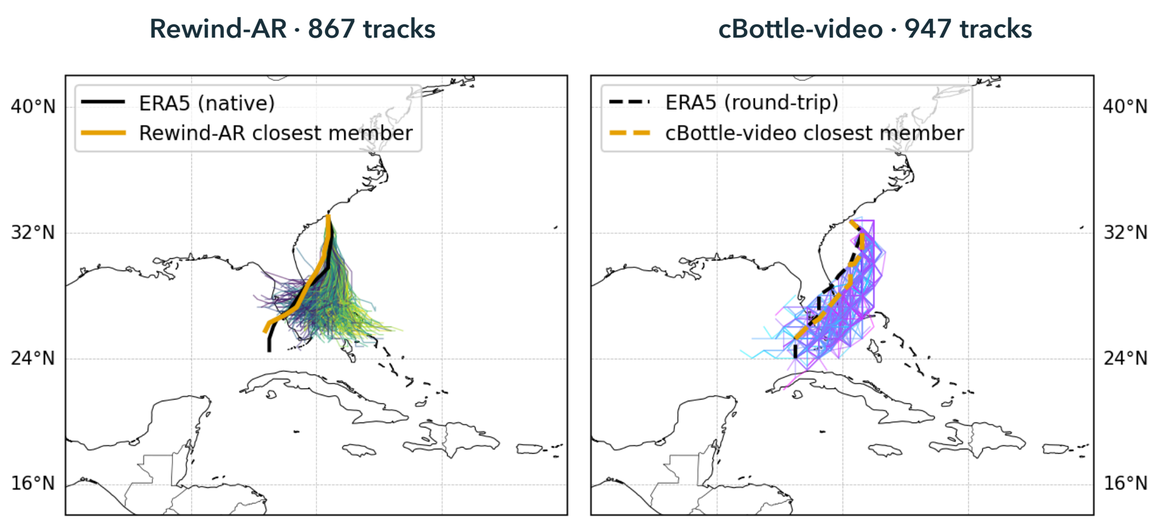
Some Rewind-AR storms start implausibly far inland



All 1,000 Rewind-AR members pass the stability screen. Among the 974 retained tracks, some start well inland over the southeastern US before reaching New Jersey. Sandy falls within both models' training periods.

HURRICANE IAN / SEPTEMBER 2022

Height patterns explain Ian's starting longitude

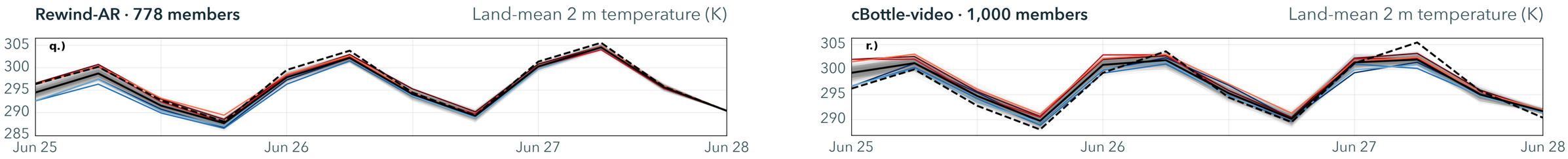


At −66 h, eight principal components of the 500 hPa height field explain **87%** of the variance in Ian's longitude for Rewind-AR and **6%** for cBottle-video (10-fold cross-validation). The height field can also reflect the storm's own position.

Track colors rank members by the first height component, from purple to yellow for Rewind-AR and cyan to magenta for cBottle-video. Black traces ERA5 and gold the closest member.

PACIFIC NORTHWEST HEATWAVE / JUNE 2021

Rewind-AR captures the rise in daily peak temperature



Rewind-AR starts 1.9 K cooler than ERA5 on average. Its daily peaks rise, but daytime temperatures average 1.0–1.6 K too cool on June 26–27. All cBottle-video members start warmer than ERA5. Their daily peaks stay nearly level. Dashed black marks ERA5. Blue and red highlight the coolest and warmest starting members. Dates are local (PDT).

Stability screening removed 22.2% of Rewind-AR heatwave members and 13.3% for Ian. Model resolution and training data also differ.

Training github.com/jerrylin96/run-it-back **Analysis** github.com/jerrylin96/rewinding-the-extremes

jlin404@bu.edu